Niveau(x) d'enseignement: 3 Numéro EDUQ : 8270387
Centre d'intérêt: 74

Genesee, Fred.

Evaluation of the 1975-1976 grade 6 French immersion class / F. Genesee, B. Stefanovic. -- [Montréal] : Commission des écoles protestantes du Grand Montréal, Instructional Services Department, 1976.
43, [1] f. : graph. ; 28 cm.

"A report submitted by the Instructional Services Department to the Protestant School Board of Greater Montreal". Bibliographie: f. 43-[44].

Cote Dewey: 372.654/G327e Cote LC: LB/1528

Descripteur(s): *Français (Langue) *Langue seconde *Classe d'immersion *Evaluation de l'apprentissage *Evaluation de programmes *Primaire (Niveau)

Présentation des résultats d'une étude visant à évaluer un programme d'immersion en langue française destiné à des élèves anglophones de sixième année. Il s'agit de comparer les habiletés en français, en anglais et en mathématiques d'élèves inscrits dans des classes d'immersion en 1975-1976 avec la situation de groupes témoins des secteurs anglophone et francophone de l'enseignement.

Au total, 189 élèves de la région de Montréal ont participé à l'enquête: 63 étaient inscrits dans des classes d'immersion alors que 66 constituent le groupe de contrôle anglophone et 60 le groupe de contrôle francophone. Sauf pour les Tests de rendement en français et en mathématiques de la CECM subis en novembre 1975, les test ont été administrés au printemps 1976: il s'agit de quatre sections du "California achievement test" et d'un exercice de composition anglaise ainsi que de tests de lecture, de compréhension auditive et de production orale en français

Les élèves des classes d'immersion ont obtenu aux tests de mathématiques et d'anglais des résultats identiques à ceux des élèves du groupe de contrôle anglophone. Ils ont également obtenu de meilleurs résultats à tous les tests de français, quoique ces résultats soient inférieurs à ceux du groupe de contrôle francophone. La performance aux tests d'aptitude linguistique académique en français (lecture, orthographe, etc.) apparaît dépendre du niveau intellectuel de l'élève, ce qui n'est pas le cas dans les tests d'habileté de communication interpersonnelle (compréhension et expression).

Type de document: r
Région(s) administrative(s): 063
Financement: Bénéficiaire - BEPGM
Organisme(s) source(s) - MEQ - DGEES
Programme(s) de subvention - PDEL
Auteur de l'analyse: ar

Evaluation of the 1975-76 grade 6 french immersion class.

Fred Genesee
B. Stefanovic

E D U Q 8270387

EVALUATION OF THE 1975- 1976 GRADE 6 FRENCH IMMERSION CLASS

F. Genesee

B. Stefanovic

A report submitted by the Instructional Services Department

to

The Protestant School Board of Greater Montreal

October 1976

ACKNOWLEDGEMENTS

We would like to express our appreciation to the school administrators, teachers and students in the PSBGM for their patience and co-operation throughout this research.

We extend a special note of thanks to the principal, staff and students of the French Control school for their generous participation in this project, and La Commission des Ecoles Catholiques de Montréal for permission to use their Tests de Rendement.

We would also like to thank Dr. George Cochrane, Director, Instructional Services Department, PSBGM, for his helpful comments on a draft of this report.

RÉSUMÉ

Les élèves fréquentant la sixième année d'immersion française (groupe d'immersion), de niveaux moyen, inférieur ou supérieur à la moyenne ont subi une batterie de tests en langue anglaise, en langue française et en mathématiques. Les tests d'anglais servaient à mesurer l'habilité de l'élève dans les domaines du vocabulaire, de la lecture, de l'orthographe et de l'écriture. Les tests de français mesuraient l'habilité en lecture, usage de la langue, compréhension auditive et expression orale. Les tests de mathématiques mesuraient l'habilité à résoudre concepts et problèmes, et ce en anglais et en français. La performance des élèves du groupe d'immersion aux tests d'anglais, de français et de mathématiques a été comparée à celle d'élèves identiques du programme normal en anglais recevant un cours de français langue seconde (groupe anglais de contrôle). La performance des élèves d'immersion aux tests de langue française a également été comparée à celle d'élèves francophones moyens étudiant dans une école française (groupe français de contrôle).

On s'est aperçu que les élèves d'immersion ont obtenu aux tests de langue anglaise des résultats identiques à ceux obtenus par les élèves du groupe anglais de contrôle. On s'est également aperçu que les élèves d'immersion ont obtenu de meilleurs résultats à tous les tests de langue française que les élèves du groupe anglais de contrôle. Les élèves d'immersion ont cependant moins bien réussi à tous les tests de langue française que les élèves du groupe français de contrôle. Les élèves d'immersion ont réussi

aux tests de mathématiques en anglais de la même façon que les élèves du groupe anglais de contrôle, mais ont mieux réussi qu'eux aux tests de mathématiques en français.

La performance aux tests d'aptitude linguistique académique en français (lecture, orthographe, etc.) dépendant du niveau intellectuel de l'élève alors qu'elle n'en dépendait pas dans les tests d'habilité de communication interpersonnelle (audition et expression). Les résultats ont été discutés en termes de facteurs contribuant au succès de l'apprentissage d'une langue seconde.

TABLE OF CONTENTS

Introduction 1

The Sample 5

The Test Battery:

(A) English Language Tests 8

(B) French Language Tests 10

(C) Mathematics 13

Results:

(A) English Language Tests 15

(B) French Language Tests 20

(C) Mathematics 29

Summary and Discussion 34

Appendices: (A) English Composition Topics 38

(B) Brief Explanation of Statistical Terminology 39

(C) Explanation of Stanine Scale 42

LIST OF TABLES

TABLE	TITLE	PAGE
1	Summary of Sample Sizes	7
2	Summary of Sample Sizes for Oral Production Evaluation	12
3	Summary of English Language Test Results	16
4	Summary of French Language Test Results	22
5	Summary of Item Analysis for the Test de Rendement en Français	23
6	Frequency Distribution of Results for The Test de Compréhension de Lecture	28
7	Summary of Mathematics Test Results	31

EVALUATION OF THE 1975-1976 GRADE 6

FRENCH IMMERSION CLASS

F. Genesee
B. Stefanovic

Instructional Services Department

The present report is the fourth in a series on the evaluation of the "Early" French Immersion program offered by the PSBGM. (See Polich, 1974, for the results from grades K to 3; Genesee, Morin & Allister, 1974, for the grade 4 results; and Genesee & Chaplin, 1975, for the grade 5 results.) The results from these evaluations had indicated that students following the French Immersion program from Kindergarten showed an initial lag in English language skills development as compared to comparable students following the regular program. However, once instruction in the English language was introduced to the Immersion students in grade 3 in the form of an English language arts course, the lag in performance disappeared and the Immersion students demonstrated the same competency in the English language (as measured by standardized tests) as did Control students in the regular program. Results from testing which did not involve standardized, normed tests also indicated that the Immersion students were developing the same native language competencies as the Control students. These results were extended at the grade 4 level to include English language writing skills.

One persistent difficulty was evident during the first four years, namely, spelling -- the Immersion students were consistently rated lower on spelling. Evidence for this was found in both the standardized achievement test results and in the results from the writing evaluation. In the grade 5 evaluations, however, and in the current evaluation, no evidence of a spelling problem was found.

Beginning with the grade 4 evaluation it was possible to assess the importance of academic achievement level upon performance in the Immersion program. At the grade 4 level a voluntary program of standardized I.Q. testing is carried out by the PSBGM. The existence of these test scores made it possible to select students from both the Immersion program and the regular English program who represented different I.Q. levels and, therefore, most likely different academic achievement levels. As a result, students in the grade 4 and grade 5 samples were selected who fell in the average, above average and below average ranges according to local PSBGM norms. It was found that children at different levels were able to benefit from the Immersion program. With respect to English language skills development, this means that, for example, the below average students in Immersion scored as well on the English language tests as did below average students in the regular program, thereby demonstrating that they were able "to hold their own" in the program and were not disadvantaged by participation in it. This was true of students at the average and above average levels as well. Similar patterns of results were found for tests which assessed academic achievement, such as mathematics. Thus, it appeared from these findings that being educated exclusively in a second language during the primary grades, K to 2, and in two languages during

the senior elementary grades, 3-6, was not an impediment to normal native language development or academic achievement among children with different levels of academic ability. This point will be discussed further in the Summary and Discussion section.

With regards to French language acquisition, it was found that the students following the Immersion program were noticeably and consistently superior to English Control students following the regular French-as-a-second language program, as was to be expected. And, although the Immersion students performed comparably well to native French-speaking Control students at the same age level, on some of the French language tests, they did not in general achieve native-like proficiency in the language. They were, nevertheless, uninhibited, creative and effective speakers of French.

An important and somewhat unexpected finding occurred with the stratification of the samples according to I.Q. level-- it was found that students from the "below average" group were indistinguishable from students in the "average" or "above average" groups when evaluated on tests of interpersonal communication skills-- listening and speaking. At the same time, the typical positive covariation between I.Q. and scholastic performance was found when tests of academic language skills in French were used-- reading, writing, spelling, and so on. These results were characteristic of students in the regular program as well as students in the Immersion program. In retrospect, this finding is not so surprising considering that most children learn to communicate orally and aurally in at least one language to a remarkably high level of sophistication before entering school and with very little apparent language teaching per se being offered in the

home. Other researchers have similarly found that I.Q. or intelligence is not necessarily the only or most important determinant of successful second language learning. (See, for example, Dockrell & Brosseau, 1967; Gardner & Lambert, 1972; and Malherbe, 1946).

The purpose of the current research was to extend our evaluation of the Immersion class to include the final year of the program.

THE SAMPLE

Students from three types of French language programs participated in the evaluation:

(1) Immersion group - students from the early French Immersion program. These students represent the third generation of students to follow this program.

(2) English Control group - students who had followed the regular English program with FSL courses. These students were selected from the same schools as were the Immersion students. These schools were Edinburgh, Meadowbrook, Roslyn and Willingdon.

(3) French Control group - native French-speaking students in a French School of La Commission des Ecoles Catholiques de Montréal.

All students were from predominantly middle class backgrounds. The students in the Immersion and English Control groups were equated on and stratified according to I.Q. There were three I.Q. groupings: above average (full scale I.Q. above 115), average (full scale I.Q. between 95 and 115), and below average (full scale I.Q. below 95). The bases of the I.Q. grouping was the students' score on the Canadian Lorge-Thorndike Tests of Intelligence (1967), which the students had written in November 1975.

There was no I.Q. information available for the French Control students. However, we were assured by the principal and teachers of these students that they were "average" students. Their performance on the 1975-76

version of the Test de Rendement en Français and en Mathématiques supports this contention: Français - X = 5.72; Mathématiques - X = 5.82. A Stanine of 5 is average.

A summary of the sample sizes and averages I.Q. scores for each group, where available, is presented in Table 1.

TABLE 1

SAMPLE SIZES

I.Q. \ Group	IMMERSION Male	IMMERSION Female	IMMERSION sub total	ENGLISH CONTROL Male	ENGLISH CONTROL Female	ENGLISH CONTROL sub total	FRENCH CONTROL Male	FRENCH CONTROL Female
Above Average	9	12	21	11	8	19	31	29
Average	14	13	27	12	14	26		
Below Average	8	7	15	10	11	21		
TOTAL	63			66			60	

AVERAGE I.Q.

	IMMERSION	ENGLISH CONTROL
Above Average	121.36	120.47
Average	107.01	106.38
Below Average	92.93	91.69

TEST BATTERY

A) English Language Tests

(i) Reading

The Reading subtest of the California Achievement Test (C.A.T.) Level 3, Form A (1970), was administered. This subtest consists of "two reference skill problems and four reading passages which typify material found in students' science, social studies, mathematics and general reading texts." Maximum possible score on this subtest is 40. Raw scores were analyzed statistically.

(ii) Vocabulary

The Vocabulary subtest of the C.A.T., as above, was administered. This subtest consists of 40 items, each consisting of a stem word in context and a list of four alternative words. The student must choose the alternative which has the best meaning for the stem word. Raw scores were analysed statistically.

(iii) Spelling

The Spelling subtest of the C.A.T., as above, was administered. It "contains 32 multiple-choice items. The student is asked to determine which word, if any, of a group of four words is misspelled." Raw scores were analyzed statistically.

(iv) Writing

Each student was asked to write a composition based on one of three topic sentences provided by the examiner. See Appendix A for a list of the topic sentences. The students completed the composition in one 45-minute period. The period was structured by the examiner so that during the first five minutes the students were asked to choose a topic sentence and to write down all the ideas they could think of that pertained to the topic. During the next five minutes the students were asked to organize these ideas into a coherent story. In the subsequent 30 minutes, they were told to write the composition. Finally, in the last five minutes they were told to review and correct their compositions.

Each composition was later scored independently by two grade 6 English teachers from the PSBGM using a blind procedure. The compositions were rated according to the following dimensions: vocabulary, spelling, sentence accuracy, sentence variety and complexity, organization, length, punctuation, originality and overall. Ratings for all of the dimensions except length, punctuation and vocabulary were made on rating scales ranging from 1 (very poor) to 5 (excellent). The ratings from both teachers for each student were averaged to give a single score for each student on each of these dimensions. These results were then subjected to analysis of variance.

The ratings for length and punctuation were dichotomous, being either "appropriate" or "inappropriate". The ratings for vocabulary were "superior", "average" or "below average". The average percentage frequency of occurrence was calculated for each rating on each of these dimensions.

B) French Language Tests

(i) Reading

A modified version of the Test de Lecture Silencieuse, Niveau 6-7 , B, (Joly, 1964) was administered. This test was designed for French Canadian schools. It consists of 23 "cloze" items; each item was made up of several sentences from which one, two or three words had been deleted. The students' task was to select from among five alternatives the one which would best fill in the blank (e.g., Quand un bébé a faim, il __1.__. 1. chante 2. marche 3. pleure 4. dort 5. écrit.). Thirty minutes were required to complete the test. Raw scores were analyzed statistically. Maximum possible score was 49.

(ii) Test de Rendement en Français, niveau 6

The grade 6 level of the Test de Rendement en Français (1974-75) was administered. This test was designed and standardized by La Commission des Ecoles Catholiques de Montréal for use in its French schools. The test consists of 40 items assessing word skills, word sense, and comprehension of text. This test was administered in November 1975 to coincide with the standardized testing program of the C.E.C.M. Raw scores were converted to stanine equivalents for statistical analysis.

(iii) Listening Comprehension

This test was designed by a research group in the PSBGM Curriculum Department. It consists of a short story entitled "Pierrot et le Cambrioleur" and 16 related, multiple-choice questions. The story was read twice, once in its entirety and a second time in shorter sections with the relevant questions

following the text. To facilitate comprehension, the text was presented by a male voice and the questions were presented by a female voice. There were four alternative answers for each question, with the final choice always being "Aucune". The test was presented by means of a tape recorder in order to standardize its administration. Raw scores were analyzed statistically.

(iv) Oral Production

A number of students were interviewed individually by a native-French-speaking person. In the first part of the interview, the student was asked to describe a cartoon picture. Subsequently, in the second part, he/she was engaged in a short conversation with the interviewer on topics related to the cartoon or the student's school experiences. The interviews were tape recorded for later analysis.

Because the analysis of oral production skills is very lengthy not all students who had taken the other French tests were evaluated in this part of the assessment. Instead, a sub-sample of the original sample was selected for evaluation. The sub-samples were selected so that, as much as possible, there would be an equal number of boys and girls in each school group. As well, for the anglophone groups, the sub-samples were selected so that there would be an equal number of students in the below average, average, and above average categories. This was not possible for the French Control group since I.Q. scores were not available. A summary of the actual sample sizes for this evaluation is presented in Table 2.

TABLE 2

Summary of Sample Sizes for
Oral Production Evaluation

	Immersion		English Control		French Control	
	Male	Female	Male	Female	Male	Female
Above Average	7	11	10	8		
Average	10	9	12	10	19	22
Below Average	7	7	9	9		

Each student's speech sample was later evaluated by two raters, working independently, on five different dimensions: compréhension, prononciation, grammaire, vocabulaire and communication. The communication dimension was intended to assess the student's willingness to speak French and his/her fluency while doing so. The ratings ranged from a maximum of 5, indicating excellent or native-like skill, to a minimum of 1, indicating very poor skill. The raters were native-French-speaking Canadians. They were not trained in linguistics. It was felt desirable to employ untrained native-French speakers since in most day-to-day situations when we speak a second language we are evaluated by non-trained people. The ratings of both raters were combined to form an average score for each dimension.

(v) Test de Compréhension de Lecture: 6e année de scolarisation

This test is designed and standardized by Le Ministère de L'Education, Gouvernement du Québec for testing of grade 6 students within the French schools of Quebec; for the most part these students speak French as a native language. The 1975-76 test consists of 40 multiple-choice questions which assess a variety of reading skills, including phoneme/grapheme correspondence, vocabulary, comprehension of text, and idiomatic usage. Raw scores on this test were converted to stanine equivalents based on the entire grade 6 student population in the French. See attached Appendix C for an explanation of stanine scores. In the PSBGM, the test was administered by classroom teachers to all grade 6 French Immersion classes. In total, 118 students wrote the test.

(C) Mathematics

The anglophone students' mathematics skills were assessed in English and French. For English, the mathematics subtests of the C.A.T., Level 3, Form A (1970), were administered; for French, the Test de Rendement en Mathématiques, niveau 6 (1975-75), was administered. Each will be described briefly.

(i) M.A.T.

The Computation (whole numbers) subtest assesses the students' ability to add, subtract, multiply and divide whole numbers. There were 12 items for each operation, making a total of 48 items in all. Raw scores were analysed statistically.

The Concepts subtest "contains 25 items which measure the students' understanding and use of mathematics concepts in a variety of contexts". Raw scores were analyzed statistically.

The Problems subtest contains 15 items and is "designed to measure the students' ability to comprehend and correctly answer a variety of word problems." Raw scores were analyzed statistically.

(ii) Test de Rendement en Mathématiques, niveau 6

This test was designed and standardized by the C.E.C.M. for use in grade 6 of its French schools for the academic year 1974-75. It consists of 40 multiple-choice questions. It must be borne in mind that performance on this test reflects, in part, differences in the mathematics curriculum of the PSBGM and that of the C.E.C.M. This test was administered in November 1975 to coincide with the standardized testing program of the C.E.C.M. Raw scores were converted to stanine equivalents, which were analyzed statistically.

The Immersion and English Control groups were administered all three sets of tests:English, French and mathematics. The French Control group was administered only the French language tests. With the exception of the Tests de Rendement, which were administered in November/December 1975, the testing took place during the spring term, 1976.

RESULTS

(A) English Language Tests

Separate three-way analyses of variance (least squares solution) were performed on the results of the following M.A.T. subtests- vocabulary, comprehension and spelling, and on the results of the following dimensions of the writing evaluation- spelling, sentence accuracy, sentence variety and complexity, organization, originality, and overall. The main variables were School Group (Immersion, English Control), I.Q. level (above average, average, below average) and Sex of Student (male, female). A summary of these results along with the corresponding F and p values is presented in Table 3. For a brief explanation of statistical terminology that will be used throughout the remainder of this report, see Appendix B.

There were no statistically significant differences between the Immersion students and the English Control students on any of the three M.A.T. language subtests - - vocabulary, comprehension and spelling. In terms of percentiles, the students scored at the 57th percentile on the vocabulary subtest, at the 50th percentile on the comprehension subtest, and at the 53rd percentile on the spelling subtest. These norms are based on an American population of 6th grade students.

As expected, there were significant differences among the I.Q. groupings on all three subtests such that the above average students scored higher than the average students, who scored higher than the below average

TABLE 3

SUMMARY OF ENGLISH LANGUAGE TESTS

(A) BY SCHOOL GROUP

	Immersion	English Control	F Ratio	df
California Achievement Test:				
Vocabulary	32.81	31.09	2.64	1,97
Spelling	22.93	24.50	2.51	1,97
Reading	28.64	26.73	2.32	1,97
Writing Skills:				
Spelling	3.22	3.27	.08	1,110
Sentence Accuracy	2.88	3.36	1.03	1,110
Sentence Complexity	2.74	2.93	.37	1,110
Organization	2.71	3.06	.75	1,110
Originality	1.92	2.17	1.34	1,110
Overall	2.64	2.80	.45	1,110
Punctuation:				
Appropriate	70.8%	76.8%		
Inappropriate	29.2%	23.2%		
Length:				
Appropriate	56.3%	45.9%		
Inappropriate	43.7%	54.1%		
Vocabulary:				
Superior	6.1%	6.4%		
Average	72.2%	71.2%		
Below average	21.7%	22.4%		

students for the most part. Interestingly, there were no significant School Group by Language Group interactions, indicating that the below average students in the Immersion program scored at the same level as their below average counterparts in the regular program; this was also true for students in the average grouping and in the above average grouping. This finding suggests that the below average students in the Immersion program did not score differentially lower than similar students in the regular program after six years plus Kindergarten of French Immersion.

There were no significant differences between the boys and the girls on any of the three subtests.

The results from the statistical testing of the writing scores revealed that there were no statistically significant differences between the Immersion and English Control Groups on any of the six dimensions of spelling, sentence accuracy, sentence variety and complexity, organization, originality and overall. The ratings ranged from a low of 1.92 for originality to a high of 3.36 for sentence accuracy, indicating generally poor writing skills, according to these two teachers.

There were no statistically significant differences on any of these six dimensions when the results were analyzed according to I.Q. level. Thus, the below average students, although rated numerically lower than the above average and average students, were not rated statistically lower than these two groups. These results were unexpected and are at variance with previous writing evaluations which have made the same comparisons at the

TABLE 3

SUMMARY OF ENGLISH LANGUAGE TESTS

(B) BY I.Q. GROUP

	Above Average	Average	Below Average	F Ratio	df
California Achievement Test:					
Vocabulary	35.12	32.76	27.97	13.46***	2,97
Spelling	25.67	22.42	23.05	4.43**	2,97
Reading	32.03	27.56	23.47	13.69***	2,97
Writing Skills:					
Spelling	3.50	3.16	3.08	1.71	2,110
Sentence Accuracy	3.07	3.42	2.86	.51	2,110
Sentence Complexity	2.82	3.01	2.68	.37	2,110
Organization	2.80	3.21	2.65	.80	2,110
Originality	1.96	2.39	1.78	1.34	2,110
Overall	2.83	2.89	2.44	1.24	2,110
Punctuation:					
Appropriate	88%	65%	71%		
Inappropriate	12%	35%	29%		
Length:					
Appropriate	57%	54%	38%		
Inappropriate	43%	46%	62%		
Vocabulary:					
Superior	11%	4%	5%		
Average	68%	74%	73%		
Below average	21%	22%	22%		

1. Scores joined by a line differ significantly from one another.

** $p < .01$; *** $p < .001$.

grades 4, 5, 6, 7 and 11 levels (Genesee & Stanley, 1976; Genesee & Chaplin, 1975). There are two possible explanations for this finding. First of all, it may be that the raters' generally stringent standards have masked differences which might exist among students with different I.Q. levels; these differences might be revealed by raters with lower standards than those of the current raters. Secondly, however, it may be that, since the majority of children who participated in this year's evaluation have already participated in two previous evaluations, in grades 4 and 5, they may be saturated by the testing program. This possibility, coupled with the unusually adverse climate in the schools during 1975-76, may have resulted in such low motivation levels that performance on this test was uniformly poor. The influence of repeated testing on the same children in educational evaluations of this sort must be assessed carefully if valid test results are to be gathered over an extended time period.

Finally, there were no statistically significant differences between the boys and girls on any of these dimensions.

Turning now to the dimensions of punctuation, length and vocabulary, we see that the English Control group, with 76.8% of their ratings on punctuation being "appropriate", were rated somewhat higher than the Immersion group, which received 70.8% "appropriate" ratings. On the length dimension, however, the Immersion group, with 56.3% of their ratings being "appropriate", were rated higher than the English Control group, which received 45.9% "appropriate" ratings. On the dimension of vocabulary, there was virtually no difference in the way the two groups were rated.

When the results are broken down by I.Q. level, there is evidence that the above average students, in general, were rated more highly than the other students.

(B) French Language Tests

Separate analyses of variance were carried out on the results from the following tests: Test de Rendement en Français, Test de Lecture, Test d'Ecoute, and the Oral Production evaluation (compréhension, prononciation, grammaire, vocabulaire, communication). Two different sets of analyses were carried out. In Analysis 1F, the results from the French Control group were included so that I.Q. was omitted as a variable. Consequently, the variables in these analyses were School Group (Immersion, English Control, French Control) and Sex of student (male, female). In Analysis 2F, the results from the French Control group were excluded so that I.Q. level could be included as a variable. Consequently, the variables in these analyses were School Group (Immersion, English Control), I.Q. level (above average, average, below average) and Sex of student (male, female). The results of Analysis 1F give us an indication of the performance of the Immersion students relative to the students in the regular French program and to native French-speaking students. The results of Analysis 2F give us an indication of the performance of students from different I.Q. levels in the Immersion and the regular programs on the various French language tests.

Analysis 1F

A summary of the average scores by School Group, along with the

corresponding F ratios, p values and d f values, is presented in Table 4A.

There were statistically significant differences among the School Groups on the Test de Rendement en Français. The results of Newman-Keuls multiple-comparisons testing indicated that each group differed significantly from the other groups; thus, the Immersion group ($\bar{X}$ = 3.25) scored higher than the English Control group ($\bar{X}$ = 1.79) but lower than the French Control group ($\bar{X}$ = 5.70). In terms of test norms, the performance of the two anglophone groups places them in the "below average" category with the Immersion group scoring one step higher than the English Control group, and the performance of the francophone group places them in the "average" caterory.

An item analysis was carried out on the Test de Rendement en Français in order to assess the students' relative performance in each skills area included in the test. Four broad skills were differentiated; they included vocabulary, comprehension of text, logical sequencing, and interpretation of text. Thirty-five of the 40 items in this test could be classified using these four categories.

As the name indicates, vocabulary items involved word skills, such as prefixing, suffixing, antonyms, synonyms, and word knowledge. There were 12 items in this category. Comprehension of text items included paragraph-length passages followed by multiple-choice questions related to the content of the passage. There were 11 items in this category. Interpretation of text included items in which the student had to discover the main idea in a passage, to determine relevant and irrelevant ideas in apassage and to supply

TABLE 4

SUMMARY OF FRENCH LANGUAGE TEST RESULTS

(A) ANALYSIS IF

	Immersion	English Control	French Control	F Ratio	df
Test de Rendement en Français	3.25	1.79	5.70	135.89***	2,177
Test de Lecture	22.46	11.73	30.93	100.13***	2,167
Test d'Ecoute	10.98	6.16	11.89	117.75***	2,167
Production Orale:					
Compréhension	4.50	3.18	5.00	113.90***	2,144
Prononciation	4.01	3.40	5.00	73.2***	2,144
Grammaire	4.13	3.39	4.88	60.84***	2,144
Vocabulaire	3.92	2.69	4.99	168.29***	2,144
Communication	3.74	2.79	4.96	212.85***	2,144

***p < .001.

relevant ideas where missing. There were 7 items in this category.

Logical sequencing included arranging sentences in chronological order and selecting appropriate concluding sentences for incomplete passages. There were 5 items in this category.

Analyses of the Immersion and English Control students' tests was done separately. The average percentage correct items for each category and school group is presented in Table 5.

TABLE 5

SUMMARY OF ITEM ANALYSES FOR THE
TEST DE RENDEMENT EN FRANÇAIS

	Immersion	English Control
Vocabulary (12 items)	34.60%	22.93%
Comprehension (11 items)	42.00%	22.74%
Interpretation (7 items)	52.80%	31.30%
Sequencing (5 items)	38.90%	22.10%

Overall, the students were most successful on the interpretation of text items with the Immersion students answering 52.80% of these items correctly and the English Control students answering 31.30% correctly. The Immersion students were then most successful on the comprehension of text

items (42.00% correct).They were somewhat less succesrful on the sequencing items (38.90% correct) and least successful with the vocabulary items (34.60% correct). The English Control students demonstrated the same relative level of success or difficulty, on all of the latter three types of items (approx. 22% correct).

Thus, the Immersion students had less difficulty with those types of items which required interpretive/ideational reading skills (comprehension and interpretation) and more difficulty with items which required more precise analytic reading skills as were involved in the sequencing and vocabulary items. Similarly, the English Control students had relatively more success on the items requiring interpretive skills. That the English Control students did not perform differentially on the remaining categories probably reflects the fact that they scored around the chance level (20%) on these items and, therefore, were finding all of these items so difficult that their performance was undifferentiated.

There was a statistically significant difference among the School Groups on the Test de Lecture -- the Immersion students ($\overline{X}$ = 22.46) scored significantly higher than the English Control students ($\overline{X}$ = 11.73) but lower than the French Control students ($\overline{X}$ = 30.93).

Similarly, there was a significant difference among the School Groups on the Test d'Ecoute - - the Immersion students ($\overline{X}$ = 10.98) scored higher than the English Control group ($\overline{X}$ = 6.16) but significantly lower than the French Control group ($\overline{X}$ = 11.89). There were statistically

significant differences among all three groups.

The girls ($\bar{X}$ = 10.01) scored significantly higher than the boys ($\bar{X}$ = 9.33) on the Test d'Ecoute. Otherwise, there were no differences in test performance between the girls and the boys.

There were statistically significant differences among the School Groups on all five dimensions of the Oral Production evaluation. The results on all five dimensions -- compréhension, prononciation, grammaire, vocabulaire and communication were the same: the Immersion group scored significantly higher than the English Control group but lower than the French Control group. Overall, the ratings of the Immersion students put them in the "very good" to excellen. category.

There were no significant differences between the girls and the boys on any of the dimensions of this evaluation.

Analysis 7F

The only results to be discussed in this section will be those pertaining to I.Q. level since the results pertaining to the school group and sex variables are essentially the same as those from Analysis 1F. A summary of the average test results by I.Q. group along with corresponding F ratios, p values and d f values, is presented in Table 48.

There were significant differences among the I.Q. groups on the Test de Rendement en Français -- the above average group ($\bar{X}$ = 2.98) scored

TABLE 4

SUMMARY OF FRENCH LANGUAGE TEST RESULTS

(B) ANALYSIS 2F

	Above Average	Average	Below Average	F Ratio	df
Test de Rendement en Français	2.98	2.36	2.22	6.88**	2,113
Test de Lecture	19.27	16.84	15.01	5.14**	2,110
Test d'Ecoute	8.93	8.71	7.98	1.72	2,110
Production Orale:					
Compréhension	4.00	3.79	3.74	1.16	2,97
Prononciation	3.81	3.73	3.58	.74	2,97
Grammaire	3.83	3.78	3.66	.39	2,97
Vocabulaire	3.39	3.37	3.12	1.39	2,97
Communication	3.35	3.26	3.15	.92	2,97

**p<.01.

significantly higher than both the average group ($\bar{X}$ = 2.36) and the below average group ($\bar{X}$ = 2.22). There was no statistically significant difference between the average and below average groups on this test.

There was also a significant I.Q. effect on the Test de Lecture -- in this case, the above average group ($\bar{X}$ = 19.27) scored higher than the below average group ($\bar{X}$ = 15.01). There were no other significant differences according to I.Q. grouping on this test.

There were no statistically significant differences according to I.Q. on the Test d'Ecoute or on any of the five dimensions of the Oral Production evaluation. Thus, as was found in previous evaluations, the below average students were rated at the same level as the average and above average students on listening comprehension and speaking skills in French.

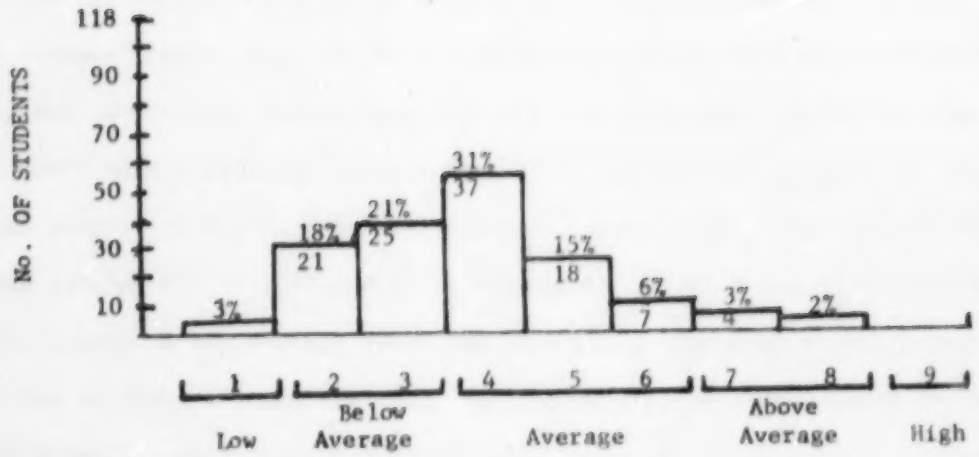
Test de Compréhension de Lecture (6e année de scolarisation)

A frequency distribution of the students' scores on this test was determined. The distribution was calculated in terms of the number of students who obtained a score at each stanine level; see Table 6.

TABLE 6

FREQUENCY DISTRIBUTION OF RESULTS FOR

THE TEST DE COMPREHENSION DE LECTURE



In total 118 students took this test. The students' scores clustered around the "average" stanine range (4-5-6) with the greatest number of scores falling in the 4th stanine (37 of 118). In terms of percentages, 52% of the students scored in the average stanine range and 31% scored at the 4th stanine level; 39% of the students scored in the "below average" range (2-3), and 5% scored in the "above average" range (6-7).

That 52% of the grade 6 French Immersion students scored in the "average" stanine range relative to provincial norms means that 52% of the Protestant School Board of Greater Montreal students scored at the same level

as the middle 54% of the 1975-76 grade 6 francophone student population throughout Quebec. For a more detailed breakdown of the percentage of cases at each stanine level for the province, see Appendix C. When compared to the provincial distribution, the distribution of Protestant School Board of Greater Montreal scores is skewed towards the "below average" end of the scale. Thus, while 39% of the Protestant School Board of Greater Montreal scores were in the "below average" range, only 19% of the provincial scores were below average. This is not surprising, considering that for the provincial sample the test was in their native language whereas for the Protestant School Board of Greater Montreal sample, it was in a second language. What is encouraging about these findings is that the largest number of Protestant School Board of Greater Montreal students scored in the average range and this group comprised slightly more than 50% of the students writing the test. This test will be administered to the 1976-77 grade 6 French Immersion group, which consists of 14 classes including a wider variety of students than in the current case.

(C) Mathematics

Separate analyses of variance (least squares solution) were carried out on the scores from the following tests --- C.A.T.: Computations, Concepts and Problem Solving and the Test de Rendement en Mathématiques (VI). In the case of the C.A.T. subtests, the main variables were School Group (Immersion, English Control), I.Q. (above average, average, below average) and Sex of Student (male, female). In the case of the Test de Rendement en Mathématiques, the main variables were School Group (Immersion, English Control, French Control) and Sex of Student (male, female). The Test de Rendement scores were analyzed a second time excluding the French Control group but including I.Q. as a variable. Consequently, the main variables in this second analysis were School Group (Immersion, English Control), I.Q. (above average, average, below average) and

A summary of the average test results, along with the corresponding F ratios, p values and df values, is presented in Table 7.

There were no significant differences between the Immersion and English Control groups on any of the three English mathematics subtests. In terms of the test norms, the groups on the average scored at the 16th percentile on the Computations subtest; and at the 72nd percentile on the Concepts and Problem Solving subtests combined.

There were significant differences among the I.Q. groups on the Concepts subtest and near significant differences ($p = .07$) on the Computations subtest. In the case of the Concepts subtest, the above average group ($\bar{X} = 22.88$) scored significantly higher than both the average ($\bar{X} = 21.31$) and the below average groups ($\bar{X} = 19.63$). There was no significant difference between the average and below average groups. In the case of the Computations subtest, the average group scored lower than the above average ($\bar{X} = 37.47$) and below average ($\bar{X} = 35.25$) groups. This latter finding is somewhat anomalous and there is no clear explanation for it.

Finally, there were no significant differences between the boys and the girls on any of the three English mathematics tests.

There was a significant difference among the school groups on the Test de Rendement en Mathématiques; the French Control group scored highest ($\bar{X} = 5.83$) followed in descending order by the Immersion group ($\bar{X} = 4.53$ and the English Control group ($\bar{X} = 3.39$). All three groups differed significantly from one another.

TABLE 7

SUMMARY OF MATHEMATICS TEST RESULTS

(A) BY SCHOOL GROUP

TEST	Immersion	English Control	French Control	F Ratio	df
C.A.T.					
Computations	35.60	34.94		.17	1,106
Concepts	21.67	20.88		1.26	1,106
Problem Solving	12.17	12.48		.15	1,106
Test de Rendement en Mathématiques	4.53	3.39	5.83	30.46***	2,177

***p< .001.

There was a significant difference according to I.Q. on the Test de Rendement en Mathématiques. The above average students ($\bar{X}$ = 4.88) scored significantly higher than the average students ($\bar{X}$ = 3.70) and the below average students ($\bar{X}$ = 3.35). There was no significant difference between the average and the below average students. Remember that in this analysis, the French Control student results have been excluded.

There was no significant difference between the boys and girls on the Test de Rendement en Mathématiques.

In summary, the Immersion students scored as well as the English Control students on the English mathematics tests and higher than the English Control students on the French mathematics test. When compared to native French-speaking students on the French mathematics test, the Immersion students scored somewhat lower although in the "average" category according to C.E.C.M. test norms. In general, performance on both the English and French mathematics tests was influenced by the studets' I.Q. such that above average students scored higher than the average and below average students.

There were no statistically significant differences between boys and girls on any of the mathematics tests.

TABLE 7

SUMMARY OF MATHEMATICS TEST RESULTS

(B) BY I.Q. GROUP

TEST	Above Average	Average	Below Average	F Ratio	df
C.A.T.					
Computations	37.47	33.09	35.25	2.66[1.]	2,106
Concepts	22.88	21.31	19.63	6.41**	2,106
Problem Solving	12.58	12.11	12.29	.12	2,106
Test de Rendement en Mathématiques	4.88	3.70	3.35	14.25***	2,113

**p< .01; ***p < .001.

1. p = .07

Summary and Discussion

The results from this year's testing indicate that

1) the Immersion students' competence in the English language was at par with that of students from the regular English program as measured by tests of vocabulary, spelling, reading and writing skills;

2) the Immersion students scored consistently higher than the English Control students on the tests of French language competence, which included reading, listening comprehension and oral production skills;

3) at the same time, however, the Immersion students scored lower than native French-speaking students on virtually all of the measures of French language skills, the one exception being the pronunciation dimension of the oral production evaluation, where there was no significant difference between the groups;

4) performance on the tests of academic language skills in French, Test de Lecture, the Test de Rendement en Français and the Test de Rendement en Mathématiques, varied positively with I.Q. level, whereas there was no similar association between I.Q. level and performance on the tests of interpersonal communication skills in French -- listening comprehension and oral production;

5) the Immersion students performed as well on the English mathematics tests as the English Control students but lower than the French Control group on the French mathematics test, although their performance on the latter test was still in the "average" range for grade 6 francophone students.

A number of comparisons can be made between the students' performance this year and last year. When interpreting these comparisons it is important to keep in mind the stressful climate in the schools during the 1975-76 academic year and the possible negative effects this might have had on the students' test performance, their attitudes towards the testing, and so on. In general, the students' performance on the 1975-76 English-language tests declined relative to their 1974-75 performance, most notably on the vocabulary subtest:

	1974-75	1975-76
Reading	57% ile	50% ile
Vocabulary	68% ile	57% ile
Spelling	56% ile	53% ile

This decline in performance characterized the immersion and the non-immersion students.

The students' performance on the two Tests de Rendement -- en Français et en Mathématiques -- did not vary appreciably from 1974-75 to 1975-76.

	Immersion		English Control	
	1974-75	1975-76	1974-75	1975-76
Test de Rendement en Français *	3.47	3.25	1.63	1.79
Test de Rendement en Mathématiques	5.36	4.52	3.04	3.44

* in stanine form

It is not possible to compare directly the results on any of the other tests owing to differences in test materials and in raters.

In general, the findings from the current evaluation corroborate those of earlier evaluations involving essentially the same students -- in grade 4 (Genesee, Morin & Allister, 1974) and in grade 5 (Genesee & Chaplin, 1975). These findings are also consistent with those from the South Shore Protestant Regional School Board, the only other school board which to date has offered French immersion from Kindergarten to the end of elementary school (Bruck, Lambert & Tucker, 1973). In addition, the results of this series of evaluations has suggested that students who place in the "below average" category on standardized group tests of intelligence are not necessarily prone to suffer differential negative effects to their native English language skills development, nor were they prone to suffer differentially academically. In other words, we have found that "below average" groups of students score as well as similar students in the regular program on tests of English language skills and on tests of academic achievement, such as mathematics. At the same time, these students seem to benefit from French immersion as much as students who place in the "average" or "above average" category in acquiring interpersonal communication skills in French.

On the other side of the coin, it is apparent that students following the French immersion program from K to grade 6, while quite competent in French, do not acquire the same level of linguistic competence as native French-speaking children of the same age. This may, in fact, be unrealistic to expect, given the limited number of French language models that the immersion program can offer. This raises the general question as to whether the level of second language competence acquired by grade 6 of the French immersion program is satisfactory.

Related to the question of the general adequacy of the level of competence is the question of the adequacy of the particular language skills that are being acquired. For example, are students who come from the French immersion program able to communicate via French comfortably and adequately in informal, day-to-day situations as well as in formal situations which tend to characterize the classroom setting? Or have the French immersion students acquired the phonological, dialectical and lexical skills that will be accepted by native French-speakers? Research conducted at Concordia University by Dr. Norman Segalowitz (1976) found that non-fluent speakers of French as a second language felt uncomfortable and reacted negatively in situations where they were required to use a casual or informal register of French. They made correspondingly negative evaluations of the French-speaking person in the situation. Segalowitz argued that these negative reactions and evaluations arose from the non-fluent French speaker's lack of competence with informal French. It is not difficult to foresee that this type of negative social situation could have a negative effect on the students' attitudes toward speaking French or their motivation to continue learning French. More detailed research would be needed to test out these hypotheses.

APPENDIX "A"

COMPOSITION TOPICS

Choose one of the following topic sentences and complete it with a story:

1) Finally my day at the OLYMPICS has arrived .....

2) With his heart pounding, he slowly approached the entrance to the dark cave ....

3) It is March 1st, 3076; my day at school has started ....

APPENDIX "B"

Analysis of variance is a mathematical procedure for testing the statistical reliability of the difference among a series of mean scores. Stated simply, an analysis of variance estimates how much of the variability in the test scores is due to the systematic effects of the experimental treatment, which in the current report is the school language program, and how much of the variability is due to non-systematic influences such as individual differences among the children. The first estimate is generally called the "treatment variance" and the second estimate is called the "error variance". The ratio of the treatment variance to the error variance is called an F ratio. If the treatment variance is larger than the error variance, the F ratio will be greater than 1. If the error variance is larger than the treatment variance, then the F ratio will be less than 1. A very small F ratio indicates that most of the variability in the scores is due to unsystematic individual differences which are unrelated to the particular treatment in question.

To decide whether a particular F ratio is "significant" or, in other words, whether there is a significant difference between the average scores, one must determine the probability (or p value) of obtaining an F ratio of a certain magnitude. The p value can be ascertained easily with the assistance of standard mathematical tables, that have already been set up, and a knowledge of the "degrees of freedom"

for your particular experiment. There is a degrees of freedom value associated with the numerator, or treatment variance, of the F ratio, and a degrees of freedom value associated with the denominator, or error variance, of the F ratio. There is no brief way of discussing the calculation or significance of degrees of freedom so that for the sake of this explanation they will not be discussed further.

According to tradition, the probability (or p value) associated with a particular F ratio must be 5 out of 100 (i.e., .05) or less ($p < .05$) before one calls the difference between the averages "significant". If the p value is greater than 5 out of 100 (i.e., $> .05$), then one calls the difference between the averages "non significant". What one is saying when one says "$p < .05$" is that the probability of getting a particular F ratio, and, therefore, a particular difference between average scores, is less than 5 in 100 by chance alone. One then goes on to assume that, since the probability is relatively low, the difference between the mean scores is probably due to the treatment and not to chance or uncontrolled influences. In other words, the treatment (or language of schooling in this report) has had a significant influence on the students' performance on the language tests. Of course, the converse rationale holds for the reverse situation where $p > .05$.

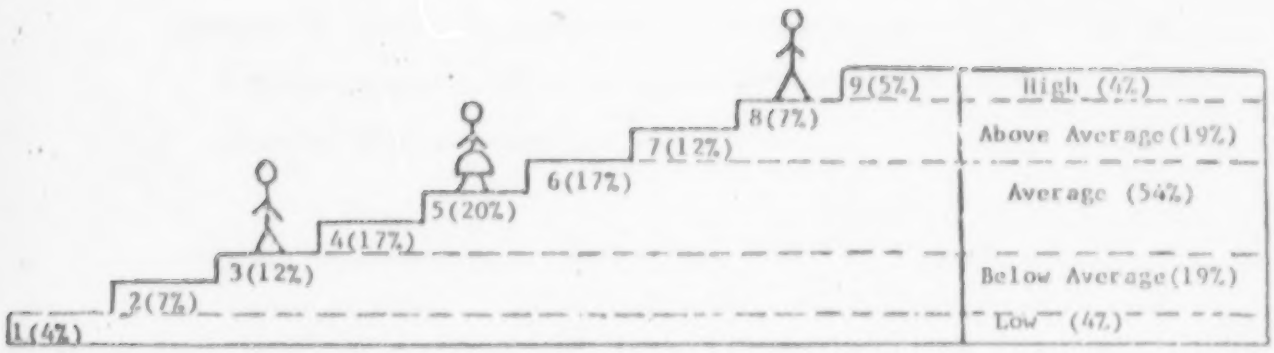
This explanation of analysis of variance, F ratios, p values and significance is of necessity brief, but hopefully will facilitate the reader's understanding of the report. More thorough and mathematical

descriptions of these concepts can be found in Statistical Analysis for Psychology and Education, 3rd ed., G.A. Ferguson, McGraw-Hill, 1971, and Statistical Principles in Experimental Design, 2nd ed., B.J. Winer, McGraw-Hill, 1971.

APPENDIX "C"

The stanine scale is a simple nine point scale of standard scores. (The word stanine was derived for STAndard NINE-point scale).

In this scale, raw scores are converted to scores which range from 1 (low) to 9 (high) with a mean of 5 and a standard deviation of 2.



REFERENCES

Bruck, M., Lambert, W.E., & Tucker, G.R. Cognitive and attitudinal consequences of bilingual schooling. The St.Lambert project through grade 6. International Journal of Psycholinguistics. In press.

Dockrell, W.B., & Brosseau, J.F. The correlates of second language learning by young children. Alberta Journal of Educational Research, 1967, 13, 295-298.

Gardner, R., & Lambert, W.E. Attitudes and motivation in second language learning. Rowley, Mass.: Newbury House, 1972.

Genesee, F., Morin, S., & Allister, T. Evaluation of the 1973-74 grade 4 French immersion class. A report submitted to The Protestant School Board of Greater Montreal, 1974.

Genesee, F., & Chaplin, S. Evaluation of the 1974-75 grade 5 French immersion class. A report submitted to The Protestant School Board of Greater Montreal, 1975.

Malherbe, E.G. The bilingual school. London: Longmans, Green & Co., 1946.

Polich, E. Report on the evaluation of the lower elementary French immersion programme through grade III, 1971-73. A report submitted to The Protestant School Board of Greater Montreal, 1974.

Segalowitz, N. Communicative incompetence and the non-fluent bilingual. Canadian Journal of Behavioral Science, 1976, 8, 122-131.